OIPE

ENTERED

RAW SEQUENCE LISTING   DATE: 05/16/2002
PATENT APPLICATION: US/09/928,847B   TIME: 14:10:20

        Input Set : A:\10022204.ST25.txt
        Output Set: N:\CRF3\05162002\I928847B.raw.

```
   3 <110> APPLICANT: Novozymes A/S
   4        Jorgensen, Steen T
   5        Rasmussen, Michael D
   6        Andersen, Jens Tonne
   7        Olsen, Carsten
   9 <120> TITLE OF INVENTION: Multiple Insertion of Genes
  11 <130> FILE REFERENCE: 10022.204-US
  13 <140> CURRENT APPLICATION NUMBER: 09/928847B
C-> 14 <141> CURRENT FILING DATE: 2002-05-06
  16 <160> NUMBER OF SEQ ID NOS: 50
  18 <170> SOFTWARE: PatentIn version 3.1
  20 <210> SEQ ID NO: 1
  21 <211> LENGTH: 30
  22 <212> TYPE: DNA
  23 <213> ORGANISM: Artificial Sequence
  25 <220> FEATURE:
  26 <223> OTHER INFORMATION: Primer
  28 <400> SEQUENCE: 1
  29 gactaagctt ctgcatagtg agagaagacg                                30
  32 <210> SEQ ID NO: 2
  33 <211> LENGTH: 67
  34 <212> TYPE: DNA
  35 <213> ORGANISM: Artificial Sequence
  37 <220> FEATURE:
  38 <223> OTHER INFORMATION: Primer
  40 <400> SEQUENCE: 2
  41 gactgaattc agatctgcgg ccgcacgcgt gtcgacagta ctgaaataga ggaaaaaata   60
  43 agtttttc                                                         67
  46 <210> SEQ ID NO: 3
  47 <211> LENGTH: 33
  48 <212> TYPE: DNA
  49 <213> ORGANISM: Artificial Sequence
  51 <220> FEATURE:
  52 <223> OTHER INFORMATION: Primer
  54 <400> SEQUENCE: 3
  55 gactgaattc cgtatccatt cctgcgatat gag                             33
  58 <210> SEQ ID NO: 4
  59 <211> LENGTH: 41
  60 <212> TYPE: DNA
  61 <213> ORGANISM: Artificial Sequence
  63 <220> FEATURE:
  64 <223> OTHER INFORMATION: Primer
  66 <400> SEQUENCE: 4
```

RAW SEQUENCE LISTING   DATE: 05/16/2002
PATENT APPLICATION: US/09/928,847B   TIME: 14:10:20

Input Set : A:\10022204.ST25.txt
Output Set: N:\CRF3\05162002\I928847B.raw

67 gactggatcc agatcttatt acaaccctga tgaatttgtc g   41
70 <210> SEQ ID NO: 5
71 <211> LENGTH: 60
72 <212> TYPE: DNA
73 <213> ORGANISM: Artificial Sequence
75 <220> FEATURE:
76 <223> OTHER INFORMATION: Primer
78 <400> SEQUENCE: 5
79 gactggatcc agatctgcta gcatcgatcc gcggctattt ccattgaaag cgattaattg   60
82 <210> SEQ ID NO: 6
83 <211> LENGTH: 31
84 <212> TYPE: DNA
85 <213> ORGANISM: Artificial Sequence
87 <220> FEATURE:
88 <223> OTHER INFORMATION: Primer
90 <400> SEQUENCE: 6
91 tatttcccga gattctgtta tcgactcgct c   31
94 <210> SEQ ID NO: 7
95 <211> LENGTH: 27
96 <212> TYPE: DNA
97 <213> ORGANISM: Artificial Sequence
99 <220> FEATURE:
100 <223> OTHER INFORMATION: Primer
102 <400> SEQUENCE: 7
103 gttttcggcc gctgtccgtt cgtcttt   27
106 <210> SEQ ID NO: 8
107 <211> LENGTH: 27
108 <212> TYPE: DNA
109 <213> ORGANISM: Artificial Sequence
111 <220> FEATURE:
112 <223> OTHER INFORMATION: Primer
114 <400> SEQUENCE: 8
115 gtgtgacgga taaggccgcc gtcattg   27
118 <210> SEQ ID NO: 9
119 <211> LENGTH: 28
120 <212> TYPE: DNA
121 <213> ORGANISM: Artificial Sequence
123 <220> FEATURE:
124 <223> OTHER INFORMATION: Primer
126 <400> SEQUENCE: 9
127 ctcttgtctc ggagcctgca ttttgggg   28
130 <210> SEQ ID NO: 10
131 <211> LENGTH: 26
132 <212> TYPE: DNA
133 <213> ORGANISM: Artificial Sequence
135 <220> FEATURE:
136 <223> OTHER INFORMATION: Primer
138 <400> SEQUENCE: 10
139 agcattattc ttcgaagtcg cattgg   26

RAW SEQUENCE LISTING   DATE: 05/16/2002
PATENT APPLICATION: US/09/928,847B   TIME: 14:10:20

Input Set : A:\10022204.ST25.txt
Output Set: N:\CRF3\05162002\I928847B.raw

```
142 <210> SEQ ID NO: 11
143 <211> LENGTH: 45
144 <212> TYPE: DNA
145 <213> ORGANISM: Artificial Sequence
147 <220> FEATURE:
148 <223> OTHER INFORMATION: Primer
150 <400> SEQUENCE: 11
151 ttaagatctt ttttatacaa ataggcttaa caataaagta aatcc                45
154 <210> SEQ ID NO: 12
155 <211> LENGTH: 3342
156 <212> TYPE: DNA
157 <213> ORGANISM: Bacillus licheniformis
159 <220> FEATURE:
160 <221> NAME/KEY: CDS
161 <222> LOCATION: (1303)..(2469)
162 <223> OTHER INFORMATION:
165 <220> FEATURE:
166 <221> NAME/KEY: misc_feature
167 <222> LOCATION: (2685)..(2685)
168 <223> OTHER INFORMATION: n denotes an undetermined nucleotide
171 <400> SEQUENCE: 12
172 gcgtaccgtt aaagtcgaac agcggtttct tcctttttac atccatggat taaaaagggg    60
174 ttgaaaaaag gtgagaaaaa gctttgtttt gcttttaacg gggctgcatg taatccttat   120
176 gctttctgcc tgcggccaaa aatcgcaaga agatgttgtg acggggctcg acaagaaggc   180
178 aaaagaatac acgtcctata aggcaaaagc gaaaatgacc attgaaacgg ggaatgaccc   240
180 gcaggagtac aacgtggaaa tctggcataa aaaaccttct ctttaccggg tctatttgga   300
182 aaacccgaaa aaagaccaga gccaggtgat cttgcgcaat gaaaacggcg tgtttgtttt   360
184 gactccgtcg ctgaataaaa gcttccgctt tcacagcgac tggcccaata acagcagcca   420
186 ggtatactta ttcgaatcgc tcgtaaagga tgtcaaaaat gatggggaag cttctttttc   480
188 cgcaaaggat tcaaaataca tttttgaaac gaaaacgaat tatcagcata atcagatgct   540
190 gccgactcag gaaatcgttt tccataaaaa gaccatggct ccttcatcgg ttaaagtgat   600
192 ggataccgac cgcaaaccga tggtaaaggt tgagtttaca agctttgaat tcgataagcc   660
194 gctcgataaa gactcttttg atgaaaagaa aaatatgacg ctgtctcaaa ttgacgtagc   720
196 gacaagcgct gacgtgtcag actctttcgc tgtcaaaacg ccgctcgatg tgcctcaggg   780
198 cgtgaaaaag cttgaagaga aagagatggc gactgaagac ggcaaacgga tcgtcatcac   840
200 atatggcggt gaaaaatcct ttacattgat tcaggaaaaa gcccgcgtcg ccaaaacatc   900
202 cacttccgta tccatgaacg gagagcccgt tgacctcggc ttcacggtcg gcgcactgac   960
204 ggataaatcg ttgtcatgga catatgacgg agtcgattac tttatctcat cagaagatct  1020
206 ttctcaagat gaacttctga tggttgcaaa aagcatgcag ggacagtctt cgaaatagac  1080
208 tgtgccgtat ccggcagcct gttttccgcc cggaagcgga aagcaggctt ttttatattt  1140
210 gcgtcgcaag cgtatgattt cgacagcttt tccgtaaaat gtataccgtg ccagcaattt  1200
212 ttctttttgtt cagggctgat gatcccgtgc aaaatttccc tttctccgaa ctttttagta  1260
214 tgatgggaag gacgagtgaa acaaggaaca ggaagtgtca tg atg agc tta aaa      1314
215                                                  Met Ser Leu Lys
216                                                    1
218 cca ttc tat aga aag aca tgg gcc gaa atc gat tta acg gct tta aaa    1362
219 Pro Phe Tyr Arg Lys Thr Trp Ala Glu Ile Asp Leu Thr Ala Leu Lys
220  5                  10                  15                  20
222 gaa aac gtc cgc aat atg aag cgg cac atc ggc gag cat gtc cgc ctg    1410
```

```
223 Glu Asn Val Arg Asn Met Lys Arg His Ile Gly Glu His Val Arg Leu
224             25              30              35
226 atg gcc gtc gtt aaa gcg aat gcc tac gga cac ggg gat gca cag gta   1458
227 Met Ala Val Val Lys Ala Asn Ala Tyr Gly His Gly Asp Ala Gln Val
228             40              45              50
230 gcg aag gcg gct ctt gca gaa ggg gcg tcc att ctt gct gtg gct tta   1506
231 Ala Lys Ala Ala Leu Ala Glu Gly Ala Ser Ile Leu Ala Val Ala Leu
232             55              60              65
234 ttg gat gaa gcg ctt tcg ctg agg gcg cag ggg att gaa gaa ccg att   1554
235 Leu Asp Glu Ala Leu Ser Leu Arg Ala Gln Gly Ile Glu Glu Pro Ile
236     70              75              80
238 ctt gtc ctc ggt gca gtg ccg acc gaa tat gca agc att gcc gcg gaa   1602
239 Leu Val Leu Gly Ala Val Pro Thr Glu Tyr Ala Ser Ile Ala Ala Glu
240 85              90              95              100
242 aag cgc att atc gtg act ggc tac tcc gtc ggc tgg ctg aaa gac gtg   1650
243 Lys Arg Ile Ile Val Thr Gly Tyr Ser Val Gly Trp Leu Lys Asp Val
244             105             110             115
246 ctc ggt ttt ctg aat gag gcc gaa gct cct ctt gaa tat cat ttg aag   1698
247 Leu Gly Phe Leu Asn Glu Ala Glu Ala Pro Leu Glu Tyr His Leu Lys
248             120             125             130
250 atc gac acg ggc atg ggc cgc ctt ggc tgc aaa acg gaa gaa gag atc   1746
251 Ile Asp Thr Gly Met Gly Arg Leu Gly Cys Lys Thr Glu Glu Glu Ile
252             135             140             145
254 aaa gaa atg atg gag atg acc gaa tcg aac gat aag ctc aat tgt acg   1794
255 Lys Glu Met Met Glu Met Thr Glu Ser Asn Asp Lys Leu Asn Cys Thr
256     150             155             160
258 ggc gtg ttc act cat ttc gcc acg gcg gac gaa aag gac acc gat tat   1842
259 Gly Val Phe Thr His Phe Ala Thr Ala Asp Glu Lys Asp Thr Asp Tyr
260 165             170             175             180
262 ttc aac atg cat ctt gac cgc ttt aaa gag ctg atc agc ccc ttc ccg   1890
263 Phe Asn Met His Leu Asp Arg Phe Lys Glu Leu Ile Ser Pro Phe Pro
264             185             190             195
266 ctt gac cgt ttg atg gtg cat tcg tca aac agc gcc gcg ggt ctg cgc   1938
267 Leu Asp Arg Leu Met Val His Ser Ser Asn Ser Ala Ala Gly Leu Arg
268             200             205             210
270 ttc agg gaa cag cta ttt aat gcc gtc cgc ttc ggc atc ggc atg tac   1986
271 Phe Arg Glu Gln Leu Phe Asn Ala Val Arg Phe Gly Ile Gly Met Tyr
272             215             220             225
274 ggt ttg gcg ccg tca acc gaa ata aaa gac gag ctg ccg ttt cgt ctg   2034
275 Gly Leu Ala Pro Ser Thr Glu Ile Lys Asp Glu Leu Pro Phe Arg Leu
276     230             235             240
278 cgg gaa gtg ttt tcg ctt cat acc gaa ctc acc cat gtg aaa aaa att   2082
279 Arg Glu Val Phe Ser Leu His Thr Glu Leu Thr His Val Lys Lys Ile
280 245             250             255             260
282 aaa aaa ggc gag agc gtc agc tac ggg gcg aca tat aca gct cag cgc   2130
283 Lys Lys Gly Glu Ser Val Ser Tyr Gly Ala Thr Tyr Thr Ala Gln Arg
284             265             270             275
286 gac gaa tgg atc ggg aca gtc ccc gtc ggg tat gcc gac gga tgg ctg   2178
287 Asp Glu Trp Ile Gly Thr Val Pro Val Gly Tyr Ala Asp Gly Trp Leu
```

RAW SEQUENCE LISTING   DATE: 05/16/2002
PATENT APPLICATION: US/09/928,847B   TIME: 14:10:20

Input Set : A:\10022204.ST25.txt
Output Set: N:\CRF3\05162002\I928847B.raw

```
288               280                 285                 290
290 agg cgc ctg gcc gga acg gaa gtg ctg atc gac gga aaa cgc caa aaa      2226
291 Arg Arg Leu Ala Gly Thr Glu Val Leu Ile Asp Gly Lys Arg Gln Lys
292         295                 300                 305            .
294 ata gca ggg aga atc tgc atg gac cag ttc atg att tcc ctt gcc gaa      2274
295 Ile Ala Gly Arg Ile Cys Met Asp Gln Phe Met Ile Ser Leu Ala Glu
296     310                 315                 320
298 gaa tac cct gtc ggc aca aag gtt acc ttg atc gga aag caa aaa gac      2322
299 Glu Tyr Pro Val Gly Thr Lys Val Thr Leu Ile Gly Lys Gln Lys Asp
300 325                 330                 335                 340
302 gaa tgg atc tca gtc gac gaa atc gcc caa aat ttg cag acg atc aat      2370
303 Glu Trp Ile Ser Val Asp Glu Ile Ala Gln Asn Leu Gln Thr Ile Asn
304                 345                 350                 355
306 tat gaa att acc tgt atg ata agt tca agg gtg ccc cgt atg ttt ttg      2418
307 Tyr Glu Ile Thr Cys Met Ile Ser Ser Arg Val Pro Arg Met Phe Leu
308             360                 365                 370
310 gaa aat ggg agt ata atg gaa ata agg aat ccg atc ttg cct gat caa      2466
311 Glu Asn Gly Ser Ile Met Glu Ile Arg Asn Pro Ile Leu Pro Asp Gln
312             375                 380                 385
314 tcc tgaaaattga tgaattagcg gaaaaacaac tttgcttgcg aaaagaataa          2519
315 Ser
318 tgatatgatt atgaatggaa tggatagagt gttgtatccg taagtttggt ggaggtgtat    2579
320 gtttttgtct gaatccagcg caacaactga aatattgatt cgcttgccag aagctttagt    2639
W--> 322 atcagaactg gatggtgtcg tcatgcgaga taaccgggag cagganatga actgatttta   2699
324 ccaagccaca aaaatgtagg aacgcgaacg caaaaaatcg acaaattcgg ggaatcgatg     2759
326 agaagcggtt atatggagat ggccaagatc caatttgaac atctcttctg aggctcaatt     2819
328 tgcagagtat gaggctgaaa acacagtaga gcgcttacta agcggatgat aatcatttga     2879
330 ttgttaaacg cggcgatgtt tattttgctg acctatctcc tgttgttggc tcagaacaag     2939
332 gcggggtgcg cccggtttta gtgattcaaa acaacatcgg caatcgcttc agcccaactg     2999
334 ctattgttgc agccataaca gcccaaatac agaaagcaaa attacctacc cacgtcgaaa     3059
336 ttgatgcgaa acgctacggt tttgaaagag actccgttat attgctcgaa caaattcgga    3119
338 cgattgacaa gcaaagatta acggacaaaa tcacccatct cgatgatgaa atgatggaaa     3179
340 aggtcaacga agccttacaa atcagtttgg cactcattga ttttttaatat tgatgaaagt    3239
342 tgctcgaggc gaaagagcaa cttttttttgt gttcaaaaat aacaatacga tataatggta    3299
344 actgttagtc ctaaaaatgt tagccagatg tagtcagggg gat                      3342
347 <210> SEQ ID NO: 13
348 <211> LENGTH: 389
349 <212> TYPE: PRT
350 <213> ORGANISM: Bacillus licheniformis
353 <400> SEQUENCE: 13
355 Met Ser Leu Lys Pro Phe Tyr Arg Lys Thr Trp Ala Glu Ile Asp Leu
356 1               5                 10                 15
359 Thr Ala Leu Lys Glu Asn Val Arg Asn Met Lys Arg His Ile Gly Glu
360             20                 25                 30
363 His Val Arg Leu Met Ala Val Val Lys Ala Asn Ala Tyr Gly His Gly
364         35                 40                 45
367 Asp Ala Gln Val Ala Lys Ala Ala Leu Ala Glu Gly Ala Ser Ile Leu
368     50                 55                 60
371 Ala Val Ala Leu Leu Asp Glu Ala Leu Ser Leu Arg Ala Gln Gly Ile
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 05/16/2002
PATENT APPLICATION: US/09/928,847B   TIME: 14:10:21

Input Set : A:\10022204.ST25.txt
Output Set: N:\CRF3\05162002\I928847B.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 2685
Seq#:49; N Pos. 669

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/928,847B**

DATE: 05/16/2002
TIME: 14:10:21

Input Set : **A:\10022204.ST25.txt**
Output Set: **N:\CRF3\05162002\I928847B.raw**

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:2639
L:1412 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:660